May 2021

The Influence of Data Mining in Increasing Benefits of Libraries in Jordanian Governmental Universities

Mohammad Niqresh
dr.naqrash@bau.edu.jo

The Influence of Data Mining in Increasing Benefits of Libraries in Jordanian Governmental Universities

Dr. Mohammad Ahmad Niqresh

Assistant Professor

Princess Alia University College- Albalqa Applied University

Jordan

Dr.naqrash@bau.edu.jo

Abstract

This current study aimed at examining the impact of data mining on increasing benefits of university library in Jordanian governmental university. Researcher adopted techniques of data mining which included (Association, Classification, Clustering, Prediction, Sequential Patterns and Decision Trees). Through employing the quantitative approach and utilizing a questionnaire as a study tool, (412) responded to an online survey which primary data later on was screened and processed using SPSS v. 27th. Results of study accepted the main hypothesis as there appeared an influence of data mining in better organization flow and accumulation of library data and better develop library's services for users. Among data mining techniques, it appeared that (Sequential pattern, decision trees and Prediction techniques) were the most influential techniques on library services followed by librarians in developing library services, this was noticed through the high correlation which connected them to the dependent variable, and the remaining variables also appeared to be positive in influence with a medium correlation. Study recommended to better data mining application by responsible parties within Jordanian universities as there appeared an acceptable level of application; however, the application isn't used to its maximum capacity.

Keywords: Libraries, Digital Library, Data Mining, Clustering, Classification, Association, Data, Information, Petabyte

Introduction

The widespread availability of information technology and its ease of availability have led to a pre-emptive increase in the volume of information that has not been witnessed before in history, which made the issue of big data on the Internet a matter of controversy, in terms of the feasibility of its existence in this random way, and when we talk about big data, we are talking about unimaginable quantities of data of multiple types and sources in the size of hundreds of terabytes or even a petabyte (a petabyte is the number one followed by 15 zeros) and IBM speaks about 2.5 quintillion bytes of data every day (the quintillion is the number one followed by 18 zeroes) (Stančin and Jović, 2019; Saliba, 2019).

In that sense, one might wonder what is the importance of this data in light of the fact that the structured information from this data does not constitute only a small part, up to 10%, compared to the unorganized information, which constitutes the rest, this has led to an increased need to develop powerful tools for analyzing data and extracting information and knowledge from it. Traditional and statistical methods cannot deal with this huge amount, so smart tools are used to process this data (Oberacher et al, 2019).

Based on the foregoing, a great need arose for a means by which to save, arrange, organize and process the huge amount of data present, which led to the emergence of what is known today as data mining (Chen et al, 2019). The idea of data mining is based on cognitive conclusions for large amounts of information based on mathematical algorithms derived from statistics, mathematics, logic, intelligence and many other smart and non-traditional sciences (Li et al, 2019). The first appearance of the principle of data mining was in the late eighties by converting data from mere unintelligible data to valuable information in order to direct benefit from it (Al-Daihani and Abrahams, 2018).

Launching from the above argument, current study seeks to examine the influence of applying data mining techniques on developing governmental university libraries services in Jordan.

Questions and Hypotheses

From the problem formulating above and launching from the main aim of study, researcher seeks to answer the following set of questions:

1. What is the reality of data mining in Jordanian universities' libraries?

2. How did data mining help in presenting better library services for users?

3. What are the attitudes of librarians regarding the employment of data mining in Jordanian university libraries?

Based on above questions researcher developed the following model from which study hypotheses were extracted:

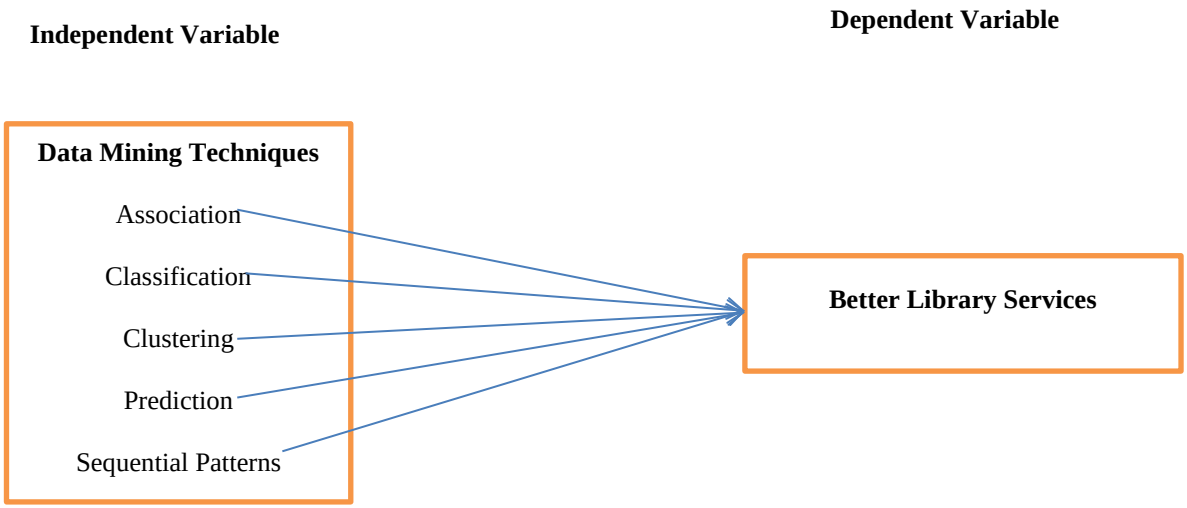

Figure 1. Study Model Sengupta (2017) and Kovacevic et al (2010)

Deeply examining the above figure, following set of hypotheses was extracted:

Main Hypothesis:

H: Data mining techniques positively influences governmental university's' library services

Sub-Hypotheses:

H1: Association techniques positively influences governmental university's' library services

H2: Classification techniques positively influences governmental university's' library services

H3: Clustering techniques positively influences governmental university's' library services

H4: Prediction techniques positively influences governmental university's' library services

H5: Sequential pattern techniques positively influences governmental university's' library services

H6: Decision trees techniques positively influences governmental university's' library services

Literature Review

Data Mining

The idea of data mining has attracted many researchers and those interested in it during the past decade and up to the present day, and many attempts have also appeared aimed at developing algorithms capable of adapting to the increasing amount of accumulated data (Yang, 2019). What was surprising is the huge amount of definitions which were presented for the concept of data mining, based on that, after going through a good number of definitions, the following were chosen:

- The process of analyzing data from different perspectives and turning it into useful information (Viloria et al, 2019)
- Create relationships between data or understand pre-existing patterns (Amin et al, 2019).
- A complete set of tools that make it easy to explore and analyze data contained in a 'Data Warehouse' or 'DataMart' database (Wang et al, 2020).
- A number of techniques that are relied upon as well as analytical tools that allow extracting important information from a large amount of data (Hartama et al, 2019)
- A technology capable of finding models and correlations between existing data, the relationship between which is usually ambiguous, and therefore data mining contributes to clarifying the relationship between this data and converting it into information (Leskovec et al, 2020).
- Data mining is complex algorithms that divide data and provide approximate possibilities of the nature of the relationship between them, that is, they represent the state of knowledge discovery in the data (Olson and Lauhoff, 2019).

Novikov (2019) indicates that data mining contributes significantly to processing large amounts of data, especially for organizations that focus mainly on consumer and customer data such as marketing organizations, finance, communications, libraries and many others.

Data Mining in Libraries

In the field of libraries benefiting - in all its forms - data mining, it was indicated by Salloum et al (2018) that the idea of data mining is of great importance in the field of libraries, as there is what is known today as "Data Mining in Libraries", which is a term that appeared for the first time in 1998, and later developed into what is known today as "Bibliomining" concerned with applying statistical tools and practices in libraries in order to classify and link data and convert it into information, in addition to assisting in decision-making or developing library services provided to beneficiaries. Zhang et al (2015) considered that data mining in the field of libraries is of great importance as it helps greatly in looking at data from a different point of view and giving it a higher value, and thus led to the emergence of many other fields of exploration, which included Web mining, Knowledge Mining, Citation. Mining, Text Mining

Hypotheses Development

According to Sengupta (2017), the idea of data mining is very important in the field of library management, it presents an approach to better decision making for librarians as it controls the amount of data flow, organize it and present it in a form of information. This information can be the source of knowledge for librarians which later can help them make informed decisions based on real and precise information. This helped in developing library services and presents a better way to make the right decision at the right time. Kovacevic et al (2010) pointed out the same idea arguing that data mining is one of the most suitable techniques that can facilitate the process of data information in a place with huge data flow like a library. Through data mining - Kovacevic et al (2010) argues – and its multiple techniques; the processing of data can be attainable. Kovacevic et al (2010) also adds that data mining is one of the tools that are suitable for digital libraries, as through its techniques like clustering, classifications and associations increased the organization level within the library and gave better chances for librarian to present good services for users.

From the perspective of Zhan (2016), the great role of data mining in improving office services was pointed out, as the researcher indicated that data mining "cleans" the existing and available data and processes it in order to ensure the extraction of high-quality information and thus achieve the library's objectives in that The link between users and knowledge is formed, in addition to providing electronic services by making the libraries' websites available to users, and therefore the amount of information increases and with this increase, the idea of data mining is classified, linked and allocated according to the information it contains and presented to the users in an understandable and unambiguous way.

Sahoo and Mishra (2015) also came up with the same idea arguing that the multiple techniques that data mining present including (Association, Classification, Clustering, Prediction, Sequential Patterns and

Decision Trees) can help the librarians follow the organization of all data the library has based on logical algorithms. Such concept paved the way for more association between data amounts, leading to better classification and reaching the state of predicting the gap, fixes it and make the right decision of either remove the service that the library present, leave it as is or develop in a certain approach.

As for Nguyen et al (2019), it was stated that the idea of sharing information has become one of the most important foundations on which knowledge is based, especially in libraries, as sharing information is an important matter in the field of knowledge and its management, and the mining of data came in order to facilitate the mechanism of organizing the data that accumulate in large quantities and convert it into Usable information and then presented to users through hardware and software in the libraries, including research, identification, selection and other activities that take place inside the libraries.

Methods

Methodological Approach

The methodological approach was seen to be as a limitation for current study. Initially, researcher aimed at examining the relationship between variables from a qualitative approach, in that sense, it was meant to meet and interview librarian face to face, ask them prompt questions and gain deeper answers which would help the main aim of study. With the breakout of COVID19, meetings were impossible as the lockdown urged the transition to online education, this mean that all students depended on the library's online websites and librarians had to perform their duties from distance. This matter urged researcher to adopt the quantitative approach in order to collect the primary data from librarian who weren't found at the premises of the libraries under examination. From that point, quantitative approach was employed in current study; numerical data were presented as the primary data which were later on processed for better results.

Study tool

The questionnaire was employed as the main data collecting tool, it was built by the researcher employing previous studies including Sengupta (2017) and Kovacevic et al (2010). The questionnaire consisted of two main parts, the first took into perspective demographics of study sample, while the other section consisted of statements related to study variables including Data Mining Techniques (Association, Classification, Clustering, Prediction, Sequential Patterns and Decision Trees) as independent variables, and Better Library Services as dependent variable.

After building the questionnaire, it was presented before a group of specialized academics in the field of data mining. They have reviewed all its statements and amended it based on the suitability and validity of each statement. In its final version, the questionnaire consisted of 35 statements based in liker 5 scale (1 strongly disagree, 2 disagree, 3 neutral, 4 agree and 5 strongly agree).

Population and Sampling

Population of study consisted of all librarians working within libraries of governmental universities in Jordan. A convenient sample of (450) was chosen in order to represent the population of study. An online questionnaire was built on Google forms and the link was distributed on the study sample through their emails, after application process, researcher was able to retrieve (412) valid questionnaires which gave a response ratio of (91.5%) as statistically accepted.

Screening and Analysis

Screening and analysis of the gathered primary data was done depending on SPSS v. 27^{th}. A reliability test was done in order to test the suitability of study tool. Cronbach's Alpha indicated Alpha value = 0.964 is greater than accepted value 0.60 that reflects the reliability of the questionnaire. The following statistical tests were also run on the data in order to gather the needed results:

- Descriptive Statistics
- Mean and Standard Deviation
- Multiple Regression
- Linear Regression

Analysis and Discussion

Demographics Results

In analyzing demographics of study respondents in table 1 below, it was revealed that majority of respondents were (males) forming (68.9%) of total sample within the age range of (28-33) forming (42.7%) of total sample and with an experience of (8-10) years holding a qualification of (BA) forming (64.1%) of total sample.

Table 1. Descriptive Statistics

	f	%
Gender		
Male	284	68.9
Female	128	31.1
Age		
22-27	72	17.5
28-33	176	42.7
34-39	112	27.2
+40	52	12.6
Education		
BA	264	64.1
MA	118	28.6
PHD	30	7.3
Experience		
2-4	120	29.1
5-7	98	23.8
8-10	130	31.6
+10	64	15.5
Total	412	100.0

Questionnaire Analysis

In analyzing responses of individuals on questionnaire statements and looking at the mean of each statements presented it was seen that all statement scored a mean that was higher than mean of scale 3.00 which is statistically considered to be positive results. The most positively answered statements appeared to be "Gaps and malfunction can be presented through prediction technique" scoring a mean of (3.54-5.00) compared to the least positively answered statement which was articulated "The decision tree is one of the most commonly used data mining techniques" and scoring a mean of (3.03/5.00). This indicated a positive attitude from respondents regarding statements of questionnaire which was later on manifested in the results of study tool reliability.

Table 2. Descriptive Statistic of Questionnaire

	Mean	Std. Deviat
Data Mining Techniques		
Association		
Patterns between library content are discovered and a relationship is built between them for easy reach	3.43	1.419
Association help in building relational patterns between library content	3.41	1.314
When association is formed, better reach and access to information appears	3.36	1.211
Librarians depend highly on data association for better content organization	3.39	1.228
With association, better services can be presented for users	3.37	1.167
Classification		
Data in the library are classified according to their field of interest	3.21	1.196
Different classification styles are presented to suit all users	3.16	1.001
Classification of library content makes information retrieval easy and fast	3.13	.948

Classification helps in storing bigger data size	3.23	.966
With classification the cloud is more attainable and organized	3.40	.997
Clustering		
Clustering gathers between related items automatically	3.27	1.028
Clustering present item each one according to their class and relationship to the main item	3.31	1.047
Clustering is one of the best mining techniques	3.33	.999
Clustering is one of the mining techniques that aid the process of library lending of books and references	3.14	1.185
With clustering, a librarian can keep books that have some kinds of similarities in one cluster or one shelf and label it with a meaningful name	3.13	1.210
Prediction		
Prediction can help in evaluating services quality of libraries	3.46	1.161
With prediction, better services can be presented for users	3.41	1.078
Prediction is a part of libraries main software to ease the process of reaching information	3.36	1.093
Prediction depends highly on search history of library users	3.32	1.044
Gaps and malfunction can be presented through prediction technique	3.54	.970
Sequential Patterns		
Sequential patterns analysis is one of data mining technique that seeks to discover or identify similar patterns	3.25	.982
Sequential patterns an highlight regular events or trends in transaction data over a period of time	3.11	1.257
With sequential patterns, a librarian can identify the most wanted services of a library	3.22	1.215
Sequential patterns can predict users' behavior	3.19	1.185
Sequential patterns can help in identifying users' frequency of library services	3.39	1.148
Decision Trees		
The decision tree is one of the most commonly used data mining techniques	3.03	1.149
The decision tree model is easy to understand for librarians	3.13	1.066
The decision tree model can give librarian multiple answers for one question	3.03	1.344
Decision tree is helpful for both librarians and library users	3.11	1.297
There is a way through decision tree to identify places of development in the library	3.08	1.252
Better Library Services		
University library is still considered to main source of information for students	3.44	1.111
With data mining, services are more organized and users are more willing to benefit from these services	3.18	1.083
Data mining managed to develop a lot of library services	3.29	.955
Data mining isn't applied fully within the university library	3.15	1.250
With better expertise, librarians can benefit greatly from data mining in libraries	3.28	1.195

Taking a general look at respondents' attitudes towards variables of study – as in table 3 below – it can be seen that respondents had positive attitudes regarding variables of study employed in the model – see figure 1 – examining the table indicated that the variable (Prediction) appeared to be the most positively answered among all with a mean of (3.41/5.00) compared to (Decision Trees) which scored the lowest mean of (3.07/5.00), however the lowest mean still scored higher than mean of scale which can be translated as statistically positive.

Table 3. Descriptive Statistic of Variables

	Mean	Std. Deviation
Association	3.3913	1.08155
Classification	3.2248	.86796

Clustering	3.2354	.92692
Prediction	3.4184	.93811
Patterns	3.2301	.99491
Decision Trees	3.0748	1.05353
Services	3.2684	.92475

Hypotheses Testing

Main Hypothesis:

H: Data mining techniques positively influences governmental university's' library services

Table 4. Testing Main Hypothesis

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.938[a]	.880	.879	.32219

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	309.428	6	51.571	496.797	.000[b]
	Residual	42.042	405	.104		
	Total	351.470	411			

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.048	.071		.682	.496
	Association	.071	.018	.083	3.857	.000
	Classification	.024	.027	.023	.879	.380
	Clustering	-.005	.025	-.005	-.203	.840
	Prediction	.206	.033	.209	6.157	.000
	Patterns	.532	.029	.572	18.504	.000
	Trees	.161	.022	.184	7.268	.000

By using Multiple regression, it was found that F= 496.797 was significant since p-value was less than 0.05, that means Data mining techniques positively influences governmental university's' library services. Also it was found that r = 0.938 reflected high level of correlation and the independent variables explain 88% of the variance in the dependent variable.

Sub-Hypotheses:

H1: Association techniques positively influences governmental university's' library services

Table 5. Testing 1st Hypothesis

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.470[a]	.221	.219	.81738

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	77.547	1	77.547	116.071	.000[b]

	Residual	273.922	410	.668		
	Total	351.470	411			

Coefficients

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.906	.133		14.369	.000
	Association	.402	.037	.470	10.774	.000

By using Linear regression, it was found that t = 10.774 was significant since p-value was less than 0.05, that means Association techniques positively influences governmental university's' library services. Also it was found that r = 0.47 reflected medium level of correlation and the independent variable explains 22.1% of the variance in the dependent variable.

H2: Classification techniques positively influences governmental university's' library services

Table 6. Testing 2nd Hypothesis

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.514[a]	.264	.262	.79437

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	92.748	1	92.748	146.979	.000[b]
	Residual	258.722	410	.631		
	Total	351.470	411			

Coefficients

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.504	.151		9.974	.000
	Classification	.547	.045	.514	12.123	.000

By using Linear regression, it was found that t= 12.123 was significant since p-value was less than 0.05, that means Classification techniques positively influences governmental university's' library services. Also it was found that r = 0.514 reflects medium level of correlation and the independent variable explains 26.4% of the variance in the dependent variable.

H3: Clustering techniques positively influences governmental university's' library services

Table 7. Testing 3rd Hypothesis

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.529[a]	.280	.278	.78579

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	98.310	1	98.310	159.215	.000[b]
	Residual	253.160	410	.617		
	Total	351.470	411			

Coefficients

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.

1	(Constant)	1.561	.141		11.095	.000
	Clustering	.528	.042	.529	12.618	.000

By using Linear regression, it was found that t= 12.618 was significant since p-value was less than 0.05, that means Clustering techniques positively influences governmental university's' library services.
Also it was found that r = 0.529 reflects medium level of correlation and the independent variable explains 28% of the variance in the dependent variable.

H4: Prediction techniques positively influences governmental university's' library services

Table 8. Testing 4th Hypothesis

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.829[a]	.687	.686	.51787

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	241.514	1	241.514	900.551	.000[b]
	Residual	109.956	410	.268		
	Total	351.470	411			

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.475	.097		4.922	.000
	Prediction	.817	.027	.829	30.009	.000

By using Linear regression, it was found that t = 30.009 was significant since p-value was less than 0.05, that means Prediction techniques positively influences governmental university's' library services.
Also it was found that r = 0.829 reflected high level of correlation and the independent variable explains 68.7% of the variance in the dependent variable.

H5: Sequential pattern techniques positively influences governmental university's' library services

Table 9. Testing 5th Hypothesis

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.906[a]	.822	.821	.39107

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	288.768	1	288.768	1888.209	.000[b]
	Residual	62.702	410	.153		
	Total	351.470	411			

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.547	.066		8.350	.000
	Patterns	.842	.019	.906	43.454	.000

By using Linear regression, it was found that t = 43.454 was significant since p-value was less than 0.05, that means Sequential pattern techniques positively influences governmental university's' library services. Also it was found that r = 0.906 reflects high level of correlation and the independent variable explains 82.2% of the variance in the dependent variable.

H6: Decision trees techniques positively influences governmental university's' library services

Table 10. Testing 6th Hypothesis

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.765[a]	.585	.584	.59636

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	205.654	1	205.654	578.251	.000[b]
	Residual	145.816	410	.356		
	Total	351.470	411			

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	1.204	.091		13.268	.000
	Trees	.671	.028	.765	24.047	.000

By using Linear regression, it was found that t = 24.047 was significant since p-value was less than 0.05, that means Decision trees techniques positively influences governmental university's' library services. Also it was found that r = 0.765 reflected high level of correlation and the independent variable explains 58.5% of the variance in the dependent variable.

Discussion

Current study aimed at examining the influence of data mining usage and employment within universities' libraries in developing library's services based on data mining techniques including (Association, Classification, Clustering, Prediction, Sequential Patterns and Decision Trees). Through depending on the quantitative approach and utilizing an online questionnaire; results of study indicated a positive relationship between data mining in libraries and the library's developed services, study was able to reach the following findings:

- Respondents had high level of awareness regarding data mining in libraries as they had the full ability to process statements of questionnaire and apply the suitable scale in their answers
- There appeared a good level of data mining application within Jordanian universities' libraries in order to run and manage the digital website of the library especially through COVID19 breakout in which individuals depended highly on the libraries' website.
- Analysis of study was able to accept the main hypothesis and there appeared a positive influence of data mining n developing services of university libraries.
- Sub-hypotheses were also accepted and all techniques of data mining had an influence on library services within university libraries. As it can be seen through hypotheses testing, the most influential data mining techniques appeared to be sequential patterns and prediction which both scored high correlation with dependent variable, it was followed by decision trees with also a strong correlation- *see tables 8, 9, 10* - , and other techniques including association, clustering and classification appeared also to be with positive influence but with medium correlation – *see tables 5, 6, 7* - .

The study demonstrated the importance of data mining in the field of libraries - of all its types - through the process of organizing data and its quantities and facilitating its transformation into information capable of serving the beneficiaries in the library. On the other hand, the study found that data mining in libraries has greatly contributed through related technologies to improving the mechanism of organizing

information and managing its flow, in addition to developing the library services that are provided. On the other hand, the study confirmed that the techniques of sequential patter, prediction, and decision trees had a great role in identifying the gaps in the library's holdings by relying on specific algorithms that may not have been known before.

These algorithms contribute to identifying deficiencies in library services, which give officials the ability and sufficient information in order to develop this data and thus develop the accompanying services. This result was consistent with the study of Nguyen et al (2019), Yang (2019) and Novikov (2019) when they confirmed that data mining in the office environment helps decision-makers avoid duplicating data and devise different cognitive patterns that enable the existence of relationships that bind topics to each other that was not clear before. In front of officials and decision-makers that there are such relationships between them), hence complementarity between subjects and acquisitions takes place. Instead of buying new holdings for a topic, this is replaced by developing new visions and divisions for existing holdings that can be utilized and presented to the beneficiaries in this topic.

On the other hand, the study found, through the above analysis, that the techniques of clustering, association and classification had a positive relationship with the development and improvement of library services by displaying and making available information in the way that the beneficiaries would like, and this information includes surveys, loan data, and access times. Other sources, especially if all this data is placed in a single database and the mechanisms of data mining are used in it and are linked to the library's holdings, activities and services. Officials have a fertile material through which they can extract information that is useful to them in all aspects of the library. This result was in agreement with Sengupta (2017) and Kovacevic et al (2010) when they indicated that data mining techniques in libraries would focus on developing library services by working on dividing non-specialized workers into more than one specialization, and consequently, there are many specializations, qualifications and degrees within one library, With the increase in the number of workers, it becomes more difficult for officials and decision-makers to grasp all their different aspects. But in the event that a single database is available that includes all the data of the library employees in terms of (their types - their social status - their residence - their specialties - their qualifications - their hobbies - their economic conditions - their experiences ... in addition to other data that can be collected on the workers), using the mechanisms of searching for Data such as classification, clustering, and association, also Zhan (2016) and Sahoo and Mishra (2015) agreed with these results arguing that we can extract strong relationships and link between these data and devise cognitive and information patterns that relate all of this to each other, and this new information was not evidently discovered or obtained except through the database of data mining, which gives the opportunity for officials and decision-makers to discover the human resources available to them. In a new and innovative way that helps them redistribute them within the various departments and activities in the library - not only according to their specialties, but according to the smart patterns discovered - which helps both the library and its employees alike.

Conclusion and Recommendations

Study was able to prove that the use of modern mechanisms and their applications in the field of libraries and information data mining) is one of the most important developments taking place in the arena now, both at the academic level and the practical level. Data Mining technology, our main topic in this article, depends mainly on statistics, especially in preparing algorithms for extracting future cognitive patterns, and from this standpoint it is not strange for libraries and information to specialize in using this technology as an approach to manage the huge amount of data that flow on libraries continuously and reorganize its appearance and nature in a way that can be transformed into information and become more useful for users. From that point, current study recommended:

- Increase the attention to be directed towards developing IT infrastructure as – with COVID 19 breakout – users are becoming more dependent on online websites of universities.
- Hold online seminars and workshops for librarians to increase their expertise and knowledge regarding data mining in the library fields
- It appeared that there is a good level of data mining application within Jordanian universities; however, the application isn't used to its maximum capacity, there should be better data mining application by responsible parties.

-

References

Al-Daihani, S. M., & Abrahams, A. (2018). Analysis of academic libraries' Facebook posts: text and data analytics. *The Journal of Academic Librarianship*, *44*(2), 216-225.

Amin, M. S., Chiam, Y. K., & Varathan, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics*, *36*, 82-93.

Chen, C., Xing, Z., & Liu, Y. (2019). What's spain's paris? mining analogical libraries from q&a discussions. *Empirical Software Engineering*, *24*(3), 1155-1194.

Hartama, D., Windarto, A. P., & Wanto, A. (2019, December). The Application of Data Mining in Determining Patterns of Interest of High School Graduates. In *Journal of Physics: Conference Series* (Vol. 1339, No. 1, p. 012042). IOP Publishing.

Kovacevic, A., Devedzic, V., & Pocajt, V. (2010). Using data mining to improve digital library services. *The Electronic Library*, *28*(6), 829-843. doi: 10.1108/02640471011093525

Leskovec, J., Rajaraman, A., & Ullman, J. D. (2020). *Mining of massive data sets*. Cambridge university press.

Li, S., Jiao, F., Zhang, Y., & Xu, X. (2019). Problems and changes in digital libraries in the age of big data from the perspective of user services. *The Journal of Academic Librarianship*, *45*(1), 22-30.

Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., García, Á. L., Heredia, I., ... & Hluchý, L. (2019). Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. *Artificial Intelligence Review*, *52*(1), 77-124.

Novikov, A. V. (2019). PyClustering: Data mining library. *Journal of Open Source Software*, *4*(36), 1230.

Oberacher, H., Reinstadler, V., Kreidl, M., Stravs, M. A., Hollender, J., & Schymanski, E. L. (2019). Annotating nontargeted LC-HRMS/MS data with two complementary tandem mass spectral libraries. *Metabolites*, *9*(1), 3.

Olson, D. L., & Lauhoff, G. (2019). Descriptive data mining. In *Descriptive Data Mining* (pp. 129-130). Springer, Singapore.

Sahoo, B., & Mishra, B. S. P. (2015). Data Mining Is A Perpetual Concept For Library And Information Science: An Estimated Overview. *International Journal of Digital Library Services*, *5*(3), 14-21.

Saliba, M. (2019). Polyelemental, multicomponent perovskite semiconductor libraries through combinatorial screening. *Advanced Energy Materials*, *9*(25), 1803754.

Salloum, S. A., Al-Emran, M., Monem, A. A., & Shaalan, K. (2018). Using text mining techniques for extracting information from research articles. In *Intelligent natural language processing: Trends and Applications* (pp. 373-397). Springer, Cham.

Sengupta, S. (2017). Applications of data mining in library and information centers: an overview, *International Journal of Current Research*, 9(1)

Stančin, I., & Jović, A. (2019, May). An overview and comparison of free Python libraries for data mining and big data analysis. In *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 977-982). IEEE.

Viloria, A., Acuña, G. C., Franco, D. J. A., Hernández-Palma, H., Fuentes, J. P., & Rambal, E. P. (2019). Integration of data mining techniques to PostgreSQL database manager system. *Procedia Computer Science*, *155*, 575-580.

Wang, S., Cao, J., & Yu, P. (2020). Deep learning for spatio-temporal data mining: A survey. *IEEE Transactions on Knowledge and Data Engineering*.

Yang, X. S. (2019). *Introduction to algorithms for data mining and machine learning*. Academic press.

Zhan, M. (2016). Exploring the feasibility of applying data mining for library reference service improvement: a case study of Turku Main Library.

Zhang, Y., Chen, M., & Liu, L. (2015, September). A review on text mining. In *2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS)* (pp. 681-685). IEEE.